(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
5. August 2004 (05.08.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/066566 A1

(51) Internationale Patentklassifikation[7]: H04L 12/56

(21) Internationales Aktenzeichen: PCT/EP2004/000213

(22) Internationales Anmeldedatum:
14. Januar 2004 (14.01.2004)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
103 01 966.9 20. Januar 2003 (20.01.2003) DE

(71) Anmelder *(für alle Bestimmungsstaaten mit Ausnahme von US)*: **SIEMENS AKTIENGESELLSCHAFT** [DE/DE]; Wittelsbacherplatz 2, 80333 München (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder *(nur für US)*: **MENTH, Michael** [DE/DE]; Hausnummer 2, 97255 Gelchsheim/Oellingen (DE). **MILBRANDT, Jens** [DE/DE]; Sieboldstr. 3 1/2, 97072 Würzburg (DE). **TRAN-GIA, Phouc** [DE/DE]; Am Hölzlein 33, 97076 Würzburg (DE).

(74) Gemeinsamer Vertreter: **SIEMENS AKTIENGESELLSCHAFT**; Posfach 22 16 34, 80506 München (DE).

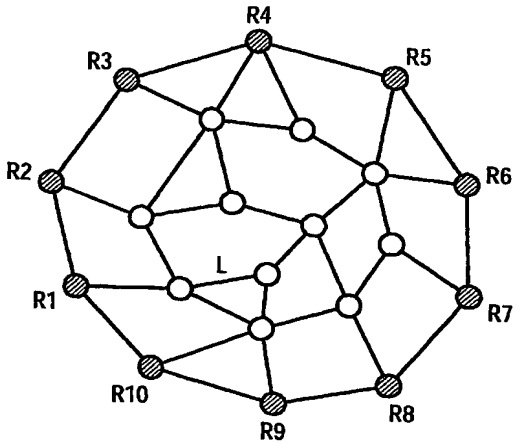
(81) Bestimmungsstaaten *(soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten *(soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart)*: ARIPO (BW,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR DETERMINING LIMITS FOR CONTROLLING TRAFFIC IN COMMUNICATION NETWORKS WITH ACCESS CONTROL

(54) Bezeichnung: VERFAHREN ZUR BESTIMMUNG VON GRENZEN FÜR EINE VERKEHRSKONTROLLE IN KOMMUNIKATIONSNETZEN MIT ZUGANGSKONTROLLE.



(57) Abstract: The invention relates to a method for determining limits for the access control of traffic that is to be transmitted via a communication network. The limits are fixed in such a way that no overload situation can occur in the network; the probability of rejection of traffic flows is, wherever possible, is independent from the point of entry into said network; and resources are used as efficiently as possible. On the basis of limits wherein no overload situation occurs, the limits for the traffic control are raised in such a way that the blocking probability for traffic transmitted between pairs of marginal modes is lowered at the same time. The lowering of said blocking probability is maintained if an overload situation were to occur in the network. For pairs (Ri, Rj) of marginal nodes contributing to the occurrence of an overload situation, the limits for traffic transmitted between the marginal nodes are fixed at a value prior to or during the overload situation. The method can be continued for the other pairs (Ri,Rj) until all limits have been set. The method results in efficient transmission of energy while maintaining quality of service parameters.

[Fortsetzung auf der nächsten Seite]

WO 2004/066566 A1

GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— *mit internationalem Recherchenbericht*

— *vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen*

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren zur Festsetzung von Grenzen für eine Zugangskontrolle von über ein Kommunikationsnetz zu übertragenden Verkehr. Die Grenzen werden so festgelegt, dass • keine Überlastsituationen im Netz auftreten können, • die Wahrscheinlichkeit für eine Abweisung von Verkehrsströmen nach Möglichkeit unabhängig von dem Eintrittspunkt in das Netz ist, und • eine möglichst effiziente Ressourcenausnutzung stattfindet. Ausgehend von Grenzen, bei denen keine Überlastsituation auftritt, werden die Grenzen für die Verkehrskontrolle so angehoben, dass die Blockerwahrscheinlichkeit simultan für zwischen Paaren von Randknoten übertragenen Verkehr abgesenkt wird. Die Absenkung wird angehalten, wenn eine Überlastsituation im Netz auftreten würde. Für Paare (Ri,Rj) von Randknoten, die einen Betrag zum Zustandekommen der Überlastsituation leisten, werden die Grenzen für den zwischen den Randknoten übertragenen Verkehr bei dem Wert festgesetzt, der bei bzw. kurz vor Auftreten der Überlastsituation erreicht ist. Das Verfahren kann für die weiteren Paare (Ri,Rj) fortgesetzt werden, bis alle Grenzen festgesetzt sind. Das Verfahren liefert einen Beitrag für eine effiziente Übertragung unter Einhaltung von Quality of Service Parametern.

Beschreibung

Verfahren zur Bestimmung von Grenzen für eine Verkehrskontrolle in Kommunikationsnetzen mit Zugangskontrolle.

Die Erfindung betrifft ein Verfahren für eine ausgewogene Festsetzung von Grenzwerten zur Verkehrsbeschränkung in einem Kommunikationsnetz mit Zugangskontrollen, wobei das Kommunikationsnetz mit Knoten und Verbindungsabschnitten gebildet ist und zumindest für einen Teil des Verkehrs, der zwischen Randknoten über das Netz übertragen werden soll, eine Zugangskontrolle mittels eines Grenzwertes vorgenommen wird.

Die Kontrolle bzw. Beschränkung des Verkehrs - Datenverkehr sowie Sprachverkehr - ist für verbindungslos operierende Kommunikationsnetze eine zentrales Problem, wenn Verkehr mit hohen Dienstgüteanforderungen, wie z.B. Sprachdaten übertragen werden sollen. Geeignete Mechanismen zur Kontrolle des Verkehrs werden derzeit von Netzwerkspezialisten, Vermittlungstechnikern und Internet-Experten untersucht.

Die derzeit möglicherweise wichtigste Entwicklung auf dem Gebiete der Netzwerke ist die Konvergenz von Sprach- und Datennetzen. In Zukunft sollen Übertragungsdienste mit verschiedensten Anforderungen über das selbe Netz übertragen wird. Dabei zeichnet sich ab, dass ein Grossteil der Kommunikation über Netze in Zukunft über verbindungslos arbeitende Datennetze, deren wichtigster Vertreter die sogenannten IP-Netze (IP: Internet Protocol) sind, vorgenommen werden wird. Die Übertragung von sogenanntem Echtzeitverkehr, z.B. Sprach- oder Videodaten über Datennetze unter Einhaltung von Dienstgütemerkmalen ist Voraussetzung für eine erfolgreiche Netzkonvergenz. Bei der Übertragung von Echtzeitverkehr über Datennetze müssen insbesondere bezüglich der Verzögerungszeiten und der Verlustrate von Datenpaketen enge Grenzen eingehalten werden.

Eine Möglichkeit für die Übertragung in Echtzeit über Datennetze unter Einhaltung von Dienstgütemerkmalen ist eine Verbindung durch das ganze Netz zu schalten, d.h. eine dem Dienst vorangehende Festlegung und Reservierung der benötigten Betriebsmittel bzw. Ressourcen. Die Bereitstellung von hinreichenden Ressourcen zu Garantie der Dienstmerkmale wird dann für jeden Verbindungsabschnitt (auch mit dem englischen Wort „Link" bezeichnet) überwacht. Technologien, die auf diese Weise vorgehen, sind beispielsweise das ATM-Verfahren (ATM: asynchronous transfer mode) oder das MPLS-Protokoll (MPLS: Multiprotocol Label Switching), welches die Festlegung von Pfaden durch IP-Netze vorsieht. Diese Verfahren haben jedoch den Nachteil hoher Komplexität und - im Vergleich zu herkömmlichen Datennetzen - geringer Flexibilität. Zustandsinformationen über die durch das Netz vermittelten Flows müssen bei den einzelnen Verbindungsabschnitten gespeichert bzw. überprüft werden.

Ein Verfahren, welches die Komplexität der verbindungsabschnittsweisen Überprüfung bzw. Kontrolle von Ressourcen vermeidet, ist das sogenannte Diff-Serv-Konzept. Dieses Konzept wird im Englischen als „stateless" bezeichnet, d.h. dass keine Zustandsinformationen über Verbindungen oder Flows entlang des Übertragungspfades vorgehalten werden muss. Stattdessen sieht das Diff-Serv-Konzept nur eine Zugangskontrolle am Netzrand vor. Bei dieser Zugangskontrolle können Pakete nach Maßgabe ihrer Dienstmerkmale verzögert, und - falls notwendig - verworfen werden. Man spricht in diesem Zusammenhang auch von traffic conditioning oder policing, von traffic shaping und traffic engineering. Das Diff-Serv-Konzept erlaubt so die Unterscheidung von verschiedenen Verkehrsklassen - man spricht hier häufig von Classes of service -, die entsprechend der Übertragungsanforderungen priorisiert oder einer geringeren Priorität behandelt werden können. Letztlich kann aber bei Datenübertragung mit Hilfe des Diff-Serv-Konzepts die Einhaltung von Dienstmerkmalen für Echtzeitverkehr nicht garantiert werden. Es stehen keine Mechanismen zur Verfügung,

den über das Netz übertragenen Echtzeitverkehr so anzupassen, dass verlässliche Aussagen über die Einhaltung der Dienstmerkmale möglich wären.

Es ist daher wünschenswert, den über ein Datennetz übertragenen Echtzeitverkehr so gut zu kontrollieren, dass einerseits Dienstmerkmale garantiert werden können und andererseits eine optimale Ressourcenausnutzung stattfindet, ohne dafür die Komplexität von durch das Netz geschalteten Verbindungen in Kauf nehmen zu müssen.

Die Erfindung hat zur Aufgabe, ein optimiertes Verfahren für die Festlegung von Grenzwerten für die Verkehrsbegrenzung in einem Kommunikationsnetz anzugeben.

Die Aufgabe wird durch ein Verfahren nach Anspruch 1 gelöst.

Erfindungsgemäß werden Grenzwerte zur Verkehrsbeschränkung in einem Kommunikationsnetz (z.B. ein IP Netz) festgelegt. Für die Datenübertragung über das Kommunikationsnetz ist zumindest für einen Teil des zu übertragenden Verkehrs - z.B. für eine oder mehrere Verkehrsklassen - vorgesehen, dass eine Zugangskontrolle vorgenommen wird, bevor Ressourcen des Netzes zur Übertragung verwendet werden. Die Zugangskontrolle findet dabei bei Randknoten des mit Knoten und Verbindungsabschnitten gebildeten Kommunikationsnetzes statt. Ein Randknoten kann dabei ein Netzzugangsknoten (auch als ingress node bezeichnet) oder ein Netzausgangsknoten (auch als egress node bezeichnet), ebenso wie ein sich im Kommunikationsnetz befindlicher End- oder Anfangsknoten einer Datenübertragung, d.h. ein Knoten des Netzes, der eine Quelle oder Senke im Hinblick auf den Verkehr darstellt, sein. Im letzteren Fall bezieht sich der Begriff „Rand" in dem Wort Randknoten nicht auf das Netz sondern auf den Übertragungspfad von Datenpaketen.

Bei der Erfindung wird von der Überlegung ausgegangen, dass eine ausgewogene Behandlung für Verkehr, der einer Zugangskontrolle unterzogen wird, bevor Netzressourcen für den Verkehr zur Verfügung gestellt werden, dann vorliegt, wenn die Wahrscheinlichkeit einer Nichtzulassung bzw. Abweisung des Verkehrs möglichst unabhängig ist von Randknoten (z.B. Netzzugangsknoten und Netzausgangsknoten) bzw. dem Übertragungspfad. In der Erfindung wird eine Mehrzahl von durch Randknoten gebildete Paare betrachtet. Ein Paar von Randknoten kann mit der Menge von möglichen, durch das Netz führenden Pfaden assoziiert werden, die zwischen den beiden Randknoten verlaufen. Bei den Paaren von Randknoten sei die Reihenfolge der Randknoten berücksichtigt, d.h. zwei Randknoten können zwei verschiedene Paare definieren, je nachdem welche Reihenfolge der beiden Randknoten betrachtet wird. Anders gesagt, bei Assoziationen von verschiedenen Pfaden mit Paaren von Randknoten ist den Pfaden eine Richtung bzw. ein Richtungssinn zugeordnet. Paare von Randknoten können beispielsweise bestehen aus einem Netzzugangsknoten und einem Netzausgangsknoten, einem Netzzugangsknoten und einem Netzknoten, der Empfänger bzw. Adressat von übertragenen Daten ist, sowie aus einem Netzknoten, der als Sender fungiert, und einem Netzausgangsknoten.

Die Wahrscheinlichkeit für die Nichtzulassung von Verkehr bzw. von Flows, die bei einem Randknoten einer Zulassungsprüfung unterworfen werden, kann mit Hilfe von Verkehrsmodellen abgeschätzt werden. Es wird davon ausgegangen, dass mit Hilfe eines Verkehrsmodells die Wahrscheinlichkeiten für Abweisung von Verkehr - im Folgenden Blockierwahrscheinlichkeit genannt - bestimmt wird. Ein derartiges Verkehrsmodell liefert beispielsweise Werte für das mittlere Verkehrsaufkommen zwischen zwei Randknoten und gibt einen Zusammenhang für die Berücksichtigung der Verkehrsschwankungen vor. Beispielsweise kann man annehmen, dass Verkehrsschwankungen einer Poisson-Verteilung gehorchen, mit der abgeschätzt werden kann, mit welcher Wahrscheinlichkeit (in unserem Fall die Blockierwahrschein-

lichkeit) der Grenzwert für die Zugangskontrolle überschritten wird. Die Blockierwahrscheinlichkeiten und die Grenzwerte für die Zugangskontrolle stehen miteinander im Zusammenhang und können in der Regel ineinander umgerechnet werden. Bei dem erfindungsgemäßen Verfahren wird für eine Mehrzahl von aus Randknoten gebildete Paare als Initialisierungsschritt die Blockierwahrscheinlichkeiten durch Festlegung der Grenzwerte für die Zugangkontrolle so eingestellt, dass sie im Wesentlichen gleich sind. Die anfänglichen Blockierungswahrscheinlichkeiten werden dabei groß genug gewählt, dass im Netz keine Überlastsituationen auftreten. Diese Festlegung entspricht einer fairen Behandlung der zwischen den Randknoten übertragenen Datenströme, insofern als dass sie mit praktisch gleicher Wahrscheinlichkeit zugelassen bzw. abgewiesen werden. Bei dieser Festlegung ist jedoch noch nicht garantiert, dass eine effiziente Nutzung der vom Netz zur Verfügung gestellten Ressourcen stattfindet. Im Hinblick auf eine effiziente Ressourcennutzung sieht das erfindungsgemäße Verfahren vor, die Blockierwahrscheinlichkeiten zu senken, d.h. die Grenzwerte für die Zulassungskontrolle entsprechend zu erhöhen, bis eine Überlastsituation auftritt. Die Absenkung der Blockierwahrscheinlichkeiten bzw. die Erhöhung der Grenzwerte für die Zulassungskontrolle wird dabei so vorgenommen, dass die Blockierungswahrscheinlichkeiten für die Paare von Randknoten im Wesentlichen gleich bleiben. Für die Paare von Randknoten, die bei dem Zustandkommen der Überlastsituation beteiligt sind, werden die Blockierwahrscheinlichkeiten im Wesentlichen auf den Wert festgesetzt, bei dem durch das Anheben der Grenzwerte die Überlastsituation verursacht wird. Beispielsweise werden die Blockierwahrscheinlichkeiten schrittweise abgesenkt und der Wert der Blockierwahrscheinlichkeiten und damit auch der Wert der entsprechenden Grenzwerte wird dann für die bei der Überlastsituation beitragenden Paare auf den Wert festgesetzt, den sie bei dem Schritt direkt vor Auftreten der Überlastsituation hatten.

Die Erfindung hat den Vorteil, dass in einem Netz ohne explizite Pfadreservierung Grenzen für die Zugangskontrolle ausgewogen und ressourceneffizient festgelegt werden können.
Entsprechend einer Weiterbildung der Erfindung werden für alle aus Randknoten gebildeten Paare der Mehrzahl von Paaren Grenzwerte festgelegt. Dabei wird für die Paare, die nicht am Zustandekommen der ersten Überlastsituation beteiligt waren, die Blockierwahrscheinlichkeit weiter simultan für alle verbleibenden Paare abgesenkt, bis eine zweite Überlastsituation auftritt. Für die am Auftreten der Überlastsituation beteiligten Paare werden die Blockierwahrscheinlichkeiten bzw. die Grenzwerte im Wesentlichen bei dem Wert eingefroren bzw. festgehalten, den sie beim Auftreten oder knapp vor dem Auftreten der Überlastsituation hatten. Dieser Schritt wird dann so lange iteriert bis für alle Paare Grenzwerte festgelegt worden sind, d.h. die Blockierwahrscheinlichkeit wird simultan für die verbleibenden Paare erhöht, bis eine Überlastsituation auftritt, bei der für die an der Überlastsituation beteiligten Paare die Blockierwahrscheinlichkeit festgehalten wird, solange bis für alle Paare die Blockierwahrscheinlichkeit feststeht.

Bei dieser Weiterbildung sind zwei Punkte zu bemerken:

1. Die Weiterbildung führt zu einer Zuordnung von Blockierwahrscheinlichkeiten bzw. Grenzwerte für alle Paare, denn eine Blockierwahrscheinlichkeit von Null für ein Paar würde bedeuten, dass man zwischen den Randpunkten des Paares unendlich viel Verkehr übertragen könnte, ohne dass es zu einer Überlastsituation käme, was für reale Netze offensichtlich nicht der Fall ist.
2. Die Blockierungswahrscheinlichkeit bzw. Grenzwerte für alle Paare der Mehrzahl von Paaren sind so festgelegt, dass eine Erniedrigung der Blockierwahrscheinlichkeit für ein beliebiges Paar aus der Mehrzahl zu einer Überlastsituation führen würde. Im diesen Sinne ist eine optimale

Ausnutzung der vom Netz der zur Verfügung stehenden Ressourcen gegeben.

Die Mehrzahl von Paaren umfasst beispielsweise sämtliche Paare aus Netzzugangsknoten und Netzausgangsknoten. In diesem Falle ist eine vollständige Kontrolle des in das Netz eintretenden und wieder aus dem Netz austretenden Verkehrs gegeben, bzw. des Verkehrs der Verkehrsklasse, die einer Zugangskontrolle unterzogen wird. Durch die Festsetzung der Grenzen bzw. der Wahl der Blockierungswahrscheinlichkeiten wird garantiert, dass keine Überlastsituation auftritt; als Konsequenz können definitive Aussagen über Dienstgütemerkmale gemacht werden. Die Festsetzung von Grenzen für die Zugangskontrolle eröffnet dann die Möglichkeit von Quality of Service-Diensten bei gleichzeitig möglichst optimaler Ausnutzung der zur Verfügung stehenden Ressourcen.

Kommunikationsnetze haben physikalische Beschränkungen für die Übertragungskapazität über die Verbindungsabschnitte oder Links, welche von dem Netz umfasst werden. Die maximale Übertragungskapazität der einzelnen Links legt eine obere Schranke für den über den jeweiligen Link übertragbaren Verkehr fest. Häufig werden Grenzen über das Verkehrsaufkommen auf den einzelnen Links niedriger als die maximale physikalische Kapazität festgesetzt, um einerseits Reserven zu haben, andererseits um Störfällen im Netz vorzubeugen. Im letzteren Fall hat man häufig die Resilience eines Netzes im Auge, d.h. die Fähigkeit auch bei Ausfällen von Netzelementen eine störungsfreie Übertragung sicherzustellen. Für das oben angesprochene Kommunikationsnetz können die Grenzen für den Verkehr auf den einzelnen Links z.B. so gewählt werden, dass der Ausfall eines (oder mehrerer) Links nicht zum Überschreiten der physikalischen Grenzen für die Kapazität der anderen Links führt, d.h. auch bei Ausfall eines Links kann der Verkehr, der einer Zulassungskontrolle unterworfen wurde, bewältigen kann. Eine Überlastsituation im Sinne des erfindungsgemäßen Verfahrens kann dann dadurch definiert werden, dass auf einen Verbin-

dungsabschnitt bzw. einen Link des Netzes die festgesetzten Grenzen für das Verkehrsaufkommen auf diesen Link überschritten werden könnten. Eine Überprüfung des Kommunikationsnetzes auf die Möglichkeit einer Überlastsituation kann z.B. mit Hilfe eines Modells für die Lastverteilung innerhalb des Netzes vorgenommen werden. Es wird beispielsweise mittels eines Simulationsprogramms kontrolliert, ob es Links des Kommunikationsnetzes gibt, für die eine im Rahmen der festgesetzten Grenzen maximale Verkehrslast zu einer Überschreitung des für die Links zulässigen Verkehrsaufkommen führen würde. Eine andere, leicht abgewandelte Definition wäre, dass die Grenze über das Verkehrsaufkommen auf den einzelnen Links mit einer hohen vorgebbaren Wahrscheinlichkeit überschritten werden würde. Die Paare von Randknoten, die zu einer durch Überschreiten eines Grenzwertes für den Verkehr auf einen Link gegebenen Überlastsituation beitragen, wären dann die, denen man Pfade zuordnen könnte, die über den Link verlaufen, der die Überlastsituation verursacht. In der Verkehrstheorie ist der Begriff Verkehrsmuster üblich, um den real (momentan) an den Eingängen des Netzes anliegenden Verkehr zu bezeichnen. Die Überprüfung auf eine Überlastsituation kann dann vorgenommen werden, indem kontrolliert wird, ob die von den Grenzen zugelassenen Verkehrsmuster in Anbetracht des im Netz vorgenommenen Routings zu einer Überlast führen oder nicht.

Im Folgenden wird der Erfindungsgegenstand im Rahmen eines Ausführungsbeispiels anhand einer Figur näher erläutert.

In der Figur ist ein aus Knoten und Links gebildetes Netz gezeigt. Dabei sind die Randknoten R1 bis R10 durch gefüllte Kreise gekennzeichnet. Die inneren Knoten sind durch nicht gefüllte Kreise dargestellt. Links sind veranschaulicht durch Verbindungen zwischen Knoten. Für das Netz können verschiedene Arten von Randbedingungen definiert werden, die eine Zulassungskontrolle am Netzrand gewährleisten. Die Art der Randbedingungen kann beispielsweise in Abhängigkeit der Topologie des Netzes gewählt werden. Die Form der Randbedingungen

entscheidet mit, bei welchen Blockierungswahrscheinlichkeiten Überlastsituation in den erfindungsgemäßen Verfahren vorkommen. Mögliche Randbedingungen sind:

1. Grenzen für den Verkehr, der zwischen zwei Randknoten übertragen wird, d.h. jeweils ein Grenzwert für ein Paar (Ri,Rj), j,i ∈ {1,..,10}, das durch zwei Randknoten gegeben ist.
2. Grenzwerte für alle Eingangs- und Ausgangsknoten. Wenn wir annehmen, dass alle Randknoten Ri, i ∈ {1,..,10} sowohl Eingangs- wie Ausgangsknoten sind, würde das 20 Grenzwerte ergeben, wobei jeweils zwei Grenzwerte, ein Eingangsgrenzwert und ein Ausgangsgrenzwert, einem Randknoten zugeordnet ist. Für einen Flow, der von dem Eingangsknoten Ri zu dem Ausgangsknoten Rj übertragen werden soll, würde dann überprüft werden, ob der Knoten die Eingangsgrenze für Ri oder die Ausgangsgrenze für Rj überschreiten würde. Bei Überschreiten wäre eine Abweisung die Folge.
3. Ein- und Ausgangsgrenzwerte wie bei 2. jedoch für alle Links des Netzes. Das heißt, man hätte für jeden Link L jeweils zwei Grenzen pro Randknoten. Für die Übertragung eines Flows vom Knoten Ri zum Knoten Rj würden die Eingangsgrenzen von Ri und die Ausgangsgrenzen von Rj geprüft werden, die sich auf Links beziehen, über die der Flow zu übertragen ist.

Im Folgenden wird der Einfachheit halber von Grenzwerten der Form 1 ausgegangen. Es sei ein Verkehrsmodell zugrunde gelegt, das die Bestimmung eines mittleren Verkehrsaufkommens zwischen zwei Randknoten Ri und Rj erlaube. Der mittlere Verkehr zwischen zwei Randknoten Ri und Rj wird der Einfachheit halber als Vij bezeichnet. Ebenso sei die Gij der Grenzwert für von dem Eingangsknoten Ri zu dem Ausgangsknoten Rj übertragenen Verkehr. Ein von Ri zu Rj zu übertragender Flow wird dann zugelassen, wenn der aggregierte Verkehr zwischen Ri und Rj nicht die Grenze Gij überschreiten würde. Dabei gelte immer j,i ∈ {1,..,10}. Das Verkehrsmodell benützt die mittleren

Verkehrswerte Vij und Annahmen über die statistischen Schwankungen, die z.B. einer Poisson-Verteilung gehorchen. Für die Initialisierung des Verfahrens werden die Grenzwerte Gij so (niedrig) festgesetzt, dass gleiche Blockierwahrscheinlichkeiten für alle Paare (Ri,Rj) bestehen und dass zudem keine Überlastsituation auftritt. Die Überprüfung auf Auftreten einer Überlastsituation kann z.B. dadurch geschehen, dass für die maximale durch die Grenzwerte zugelassene Verkehrslast unter Einbeziehung des Routings innerhalb des Netzes die Verkehrslast der einzelnen Links bestimmt und mit den Grenzen bzw. Kapazitäten der Links verglichen wird. Erfindungsgemäß werden die Blockierwahrscheinlichkeiten im selben prozentualen Verhältnis abgesenkt und die Grenzwerte Gij entsprechend erhöht. Dabei wird Hilfe des Verkehrsmodells für einen Satz von reduzierten, gleichen Werten der Blockierwahrscheinlichkeit ein korrespondierender Satz Grenzwerte Gij (analytisch oder numerisch) ermittelt, die - im Rahmen des Verkehrsmodells - eine Nichtzulassung mit der reduzierten Blockierwahrscheinlichkeit für sämtliche Paare (Ri,Rj) festsetzen. Es folgt eine Überprüfung auf Überlast. Falls keine Überlast auftritt, wird die Blockierwahrscheinlichkeit weiter simultan für alle Kommunikationsbeziehungen erniedrigt. Dies kann beispielsweise durch schrittweise Erniedrigung um 10 % des Ausgangswertes geschehen. Bei einem, z.B. dem fünften Schritt, trete eine Bottleneck bzw. eine Überlastsituation auf dem Link L auf, d.h. die Grenze für die Kapazität auf diesen Link würde durch die Wahl der Grenze bei Schritt 5 überschritten werden. Zu dieser Überlastsituation tragen beispielsweise die Paare (R1,R2), (R2,R1), (R1,R3), (R3,R1), (R1,R4) und (R4,R1) bei. Für diese Paare werden dann die Grenzen Gij bzw. die Blockierwahrscheinlichkeit auf ihren Wert bei Schritt 4 festgesetzt. Im Folgenden wird das Verfahren für die verbliebenen Paare (Ri,Rj) fortgesetzt, bis Grenzen Gij für alle Paare (Ri,Rj) festgelegt sind.

Patentansprüche

1. Verfahren für die Festsetzung von Grenzwerten zur Verkehrsbeschränkung in einem Kommunikationsnetz mit Zugangskontrollen, wobei das Kommunikationsnetz mit Knoten und Verbindungsabschnitten gebildet ist und zumindest für einen Teil des Verkehrs, der zwischen Randknoten über das Netz übertragen werden soll, eine Zugangskontrolle mittels eines Grenzwertes vorgenommen wird,
bei dem
- für eine Mehrzahl von durch Randknoten gebildeten Paaren (Ri,Rj) die Grenzwerte für die Zugangskontrolle so festgesetzt werden,
-- dass die Wahrscheinlichkeit für eine Nichtzulassung von einer Zulassungskontrolle unterzogenen Verkehrs im wesentlichen gleich ist für die Übertragung zwischen den zwei Randknoten eines jeden Paars, und
-- dass Überlastsituationen nicht auftreten,
- die jeweiligen Grenzwerte solange angehoben werden, bis eine Überlastsituation auftritt, wobei die Anhebung so vorgenommen wird,
dass die Wahrscheinlichkeit für eine Nichtzulassung von Verkehr für die einzelnen Paare (Ri,Rj) im wesentlichen gleich bleibt, und
- die Grenzwerte für die Zugangskontrolle für die Übertragung zwischen Paaren von Randknoten, bei denen der zwischen den Randknoten übertragene Verkehr zur Überlastsituation beiträgt, auf im wesentlichen den Wert festgesetzt werden, bei dem durch das Anheben der Grenzwerte die Überlastsituation verursacht wird.

2. Verfahren nach Anspruch 1,
dadurch gekennzeichnet,
dass bis zur Festsetzung aller Grenzwerte für die Zugangskontrolle für die Übertragung zwischen den Randknoten der einzelnen Paare (Ri,Rj)

- noch nicht festgesetzte Grenzwerte für die Zugangskontrolle weiter angehoben werden, bis eine Überlastsituation auftritt, wobei die Anhebung so vorgenommen wird, dass die Wahrscheinlichkeit für eine Nichtzulassung von Verkehr zwischen Paaren (Ri,Rj), für die noch keine Grenzwerte festgesetzt wurden, im wesentlichen gleich bleibt, und
- die Grenzwerte für die Zugangskontrolle für die Übertragung zwischen Paaren (Ri,Rj) von Randknoten, bei denen der zwischen den Randknoten übertragene Verkehr zur Überlastsituation beiträgt, auf im wesentlichen den Wert festgesetzt werden, bei dem durch das Anheben der Grenzwerte die Überlastsituation verursacht wird.

3. Verfahren nach Anspruch 1 oder 2,
dadurch gekennzeichnet,
dass Randknoten durch Netzzugangsknoten und Netzausgangsknoten gegeben sind.

4. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass die Randknoten Knoten des Netzes umfassen, die Quellen oder Senken von Verkehr darstellen.

5. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass die Mehrzahl von durch Randknoten gebildete Paare (Ri,Rj) alle Paare (Ri,Rj) aus jeweils einen Netzzugangsknoten und einen Netzausgangsknoten umfassen.

6. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass für den gesamten Verkehr einer Verkehrsklasse Zugangskontrollen durchgeführt werden.

7. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,

dass die Zugangskontrollen die Zulassung oder Abweisung einzelner Flows betreffen.

8. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass eine Überlastsituation dadurch gegeben ist, dass in einem Szenario hoher Verkehrsbelastung, bei dem die Grenzwerte für die Zugangskontrollen noch eingehalten werden, auf einem Verbindungsabschnitt ein Schwellenwert für den über den Verbindungsabschnitt übertragenen Verkehr überschritten wird.

9. Verfahren nach Anspruch 8, dadurch gekennzeichnet, dass den Verbindungsabschnitten des Kommunikationsnetzes Schwellenwerte für den über den jeweiligen Verbindungsabschnitt übertragenen Verkehr so zugeordnet sind, dass bei Ausfall eines oder mehrerer Verbindungsabschnitte der im Rahmen der Zulassungskontrollen zugelassene Verkehr keine Überlast darstellt.

10. Netzknoten mit Mitteln zur Durchführung eines Verfahrens nach einem der Ansprüche 1 bis 9.

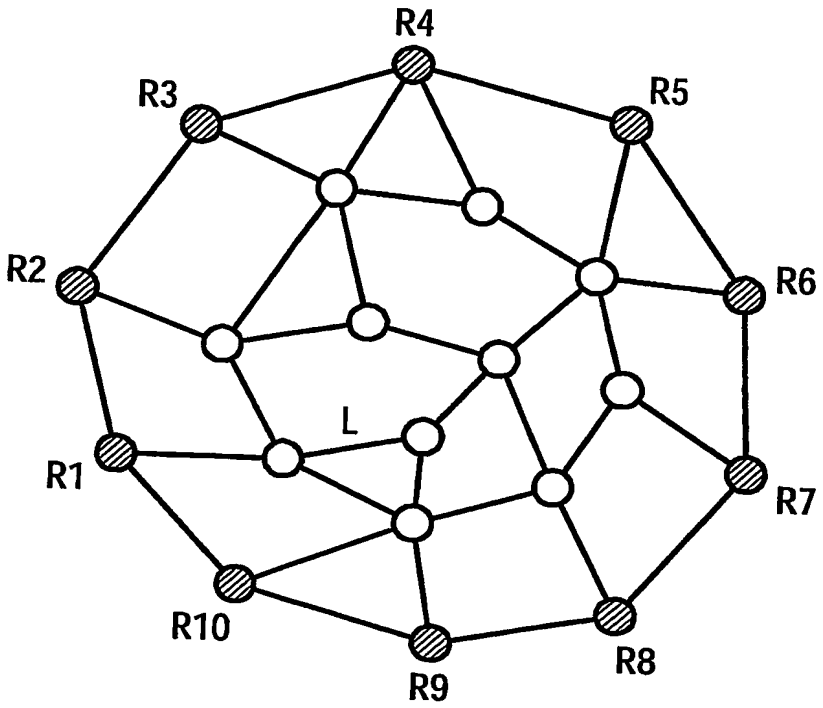


FIG. 1

INTERNATIONAL SEARCH REPORT

International Application No
PC[redacted]2004/000213

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04L12/56

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 01/28167 A (ERICSSON TELEFON AB L M) 19 April 2001 (2001-04-19) page 8, line 21 - page 9, line 15; figure 1 page 16, line 11 - page 20, line 25 -/--	1,2,6,7, 10

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 June 2004	07/07/2004

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Perrier, S

4

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

[redacted]

International Application No
[redacted]P2004/000213

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	RANDHAWA T S ET AL: "PERFORMANCE EVALUATION OF BANDWIDTH PARTITIONING IN BROADBAND NETWORKS" PROCEEDINGS OF THE IEEE CONFERENCE 2000 ON HIGH PERFORMANCE SWITCHING AND ROUTING. HEIDELBERG, GERMANY, JUNE, 26 - 29, 2000, PROCEEDINGS OF THE IEEE CONFERENCE ON HIGH PERFORMANCE SWITCHING AND ROUTING, NEW YORK, NY : IEEE, US, 26 June 2000 (2000-06-26), pages 411-418, XP001075729 ISBN: 0-7803-5884-8 page 413, left-hand column, line 30 - page 417, right-hand column, line 16	1-10
A	DE NITTO PERSONE V., GRASSI V.: "Optimal access control for integrated services wireless networks" COMPUTER COMMUNICATIONS - ELSEVIER, 'Online! 25 November 1998 (1998-11-25), pages 1559-1570, XP002284845 NETHERLANDS Retrieved from the Internet: URL:http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TYP-3VHWJ36-6-3Y&_cdi=5624&_orig=search&_coverDate=11%2F25%2F1998&_sk=999789982&view=c&wchp=dGLbVlz-zSkzV&_acct=C000049880&_version=1&_userid=987766&md5=daf08459f2b4fa98cca24eba9b0760ba&ie=f.pdf> 'retrieved on 2004-06-16! page 1560, right-hand column, line 52 - page 1561, right-hand column, line 39	1,2,6,7, 10

Form PCT/ISA/210 (continuation of second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International Application No
P2004/000213

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0128167 A	19-04-2001	AU 7696300 A	23-04-2001
		CN 1379940 T	13-11-2002
		EP 1232609 A1	21-08-2002
		JP 2003511976 T	25-03-2003
		WO 0128167 A1	19-04-2001
		SE 517146 C2	23-04-2002
		SE 0001513 A	15-04-2001

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
[redacted] 2004/000213

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 H04L12/56

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 H04L

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)
EPO-Internal, WPI Data, PAJ, INSPEC

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	WO 01/28167 A (ERICSSON TELEFON AB L M) 19. April 2001 (2001-04-19) Seite 8, Zeile 21 - Seite 9, Zeile 15; Abbildung 1 Seite 16, Zeile 11 - Seite 20, Zeile 25 -/--	1,2,6,7, 10

[X] Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

[X] Siehe Anhang Patentfamilie

° Besondere Kategorien von angegebenen Veröffentlichungen :
"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist
"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden
"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist
"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der Internationalen Recherche	Absendedatum des Internationalen Recherchenberichts
24. Juni 2004	07/07/2004
Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bevollmächtigter Bediensteter Perrier, S

4

Formblatt PCT/ISA/210 (Blatt 2) (Januar 2004)

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
[redacted]P2004/000213

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	RANDHAWA T S ET AL: "PERFORMANCE EVALUATION OF BANDWIDTH PARTITIONING IN BROADBAND NETWORKS" PROCEEDINGS OF THE IEEE CONFERENCE 2000 ON HIGH PERFORMANCE SWITCHING AND ROUTING. HEIDELBERG, GERMANY, JUNE, 26 - 29, 2000, PROCEEDINGS OF THE IEEE CONFERENCE ON HIGH PERFORMANCE SWITCHING AND ROUTING, NEW YORK, NY : IEEE, US, 26. Juni 2000 (2000-06-26), Seiten 411-418, XP001075729 ISBN: 0-7803-5884-8 Seite 413, linke Spalte, Zeile 30 - Seite 417, rechte Spalte, Zeile 16	1-10
A	DE NITTO PERSONE V., GRASSI V.: "Optimal access control for integrated services wireless networks" COMPUTER COMMUNICATIONS - ELSEVIER, 'Online! 25. November 1998 (1998-11-25), Seiten 1559-1570, XP002284845 NETHERLANDS Gefunden im Internet: URL:http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TYP-3VHWJ36-6-3Y&_cdi=5624&_orig=search&_coverDate=11%2F25%2F1998&_sk=999789982&view=c&wchp=dGLbVlz-zSkzV&_acct=C000049880&_version=1&_userid=987766&md5=daf08459f2b4fa98cca24eba9b0760ba&ie=f.pdf> 'gefunden am 2004-06-16! Seite 1560, rechte Spalte, Zeile 52 - Seite 1561, rechte Spalte, Zeile 39	1,2,6,7, 10

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
[redacted]2004/000213

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 0128167 A	19-04-2001	AU 7696300 A	23-04-2001
		CN 1379940 T	13-11-2002
		EP 1232609 A1	21-08-2002
		JP 2003511976 T	25-03-2003
		WO 0128167 A1	19-04-2001
		SE 517146 C2	23-04-2002
		SE 0001513 A	15-04-2001

Description

Method for determining limits for controlling traffic in communication networks with access control.

The invention relates to a method for a balanced determination of values for limiting traffic in a communication network with access controls, with the communication network being formed by nodes and connection links and with access being controlled by means of a limit value for at least a part of the traffic which is to be transmitted between marginal nodes over the network.

The control or limiting of the traffic - both data traffic and voice traffic - is a central problem for communication networks which use connectionless operation where traffic is to be transmitted with high quality-of-service requirements, such as voice data transmission for example. Suitable mechanisms for checking the traffic are currently being investigated by network specialists, telecommunications engineers and Internet experts.

Possibly the most important current development in the network area is the convergence of voice and data networks In the future transmission services with a very wide diversity of requirements will be transmitted over the same network. The feature which marks out such developments is that a large part of the communication over networks in the future will be via networks which operate in connectionless mode, the most important representative of which is what is known as the IP (IP: Internet Protocol) network. The transmission of what is

referred to as realtime traffic, e.g. voice or video data over data networks while preserving quality-of-service features is the prerequisite for successful network convergence. For the transmission of realtime traffic over data networks in particular narrow limits have to be adhered to as regards delay times and the packet loss rate of data packets.

One possibility for transmission in real time over data networks while maintaining quality-of-service features is to switch a connection through the entire network, i.e. to define and reserve in advance the operating means or resources required for the service. The provision of sufficient resources to guarantee the service features is then monitored for each connection section (also known as a „link"). Technologies which operate in this way are for example ATM (ATM: Asynchronous Transfer Method) or the MPLS (MPLS: Multiprotocol Label Switching) protocol which provides for the definition of paths through IP networks. The disadvantage of these methods however is their great complexity and - in comparison to conventional data networks - lower flexibility Status information about the flows switched through the network must be stored or checked for the individual links.

A method which avoids the complexity of link-by-link checking or control of resources is what is known as the Diff-Serv concept. This concept is referred to as „stateless" to indicate that no status information about data connections or flows along the transmission path has to be maintained. Despite this the Diff-Serv concept only provides for access control at the margins of the network. With this access control packets can be delayed in accordance with their service features, and - if necessary - discarded. This is also

described as traffic conditioning or policing, traffic shaping and traffic engineering. The Diff-Serv concept thus allows a distinction to be made between different traffic classes - frequently called classes of service - which can be prioritized in accordance with the transmission requirements and/or handled with a lower priority. Lastly with data transmission with the aid of the Diff-Serv concept it is not possible to guarantee that service features are maintained for realtime traffic. There are no mechanisms available to adapt the realtime traffic transmitted over the network so that reliable statements about the maintenance of the service features would be possible.

It is thus desirable for the control of the realtime traffic transmitted over the data network to be good enough that on the one hand service features can be guaranteed and on the other hand optimum use is made of resources, without having to take account of the complexity of connections switched through the network.

The object of the invention is to specify an optimized method for the definition of limit values for traffic restriction in a communication network.

The object is achieved by a method according to claim 1.

In accordance with the invention limit values are defined for limiting traffic in a communication network (e.g. an IP network). For data transmission over the communication network there is provision for access control to be undertaken for at least a part of the traffic to be transmitted - e.g. for one

or more classes of service before resources of the network are used for transmission. Access is controlled in this case at marginal nodes of the communication network formed by nodes and links. A marginal node in this case can be a network access node (also known as an ingress node) or a network output node (also known as an egress node), as well as an end or start note of a data transmission located in the communication network, i.e. a node of the network which represents a source or sink as regards the traffic. In the latter case the term „marginal" in the word marginal node does not refer to the network but to the transmission path of data packets.

The starting point for the invention is the consideration that balanced traffic handling which is subject to access control, before network resources are made available for the traffic, is present if the likelihood of a non-approval or rejection of the traffic is as independent as possible from the marginal nodes (e.g. ingress nodes and egress nodes) or the transmission path. The invention will look at a plurality of pairs formed by the marginal nodes. A pair of marginal nodes can be associated with the set of possible paths leading through the network which run between the two marginal nodes. With the pairs of marginal nodes the sequence of the marginal nodes is taken into account, i.e. two marginal nodes can define two different pairs depending on how the sequence of the two marginal nodes is looked at. In other words, for association of different paths with pairs of marginal nodes the paths are assigned a direction or a direction sense. Pairs of marginal nodes can for example consist of a an ingress node and an egress node, of an ingress node and a network node which can be receiver or addressee of transmitted data, as

well as a network node which functions as a transmitter, and an egress node.

The probability of non-approval of traffic or of flows which are subject to an approval check at an marginal node can be estimated using traffic models. The invention starts from the assumption that with the aid of a traffic model the probability of rejection of traffic - referred to below as blocking probability - will be determined. This type of traffic model typically delivers values for the average traffic intensity between two marginal nodes and specifies a relationship for taking the traffic fluctuations into account. For example it can be assumed that traffic fluctuations belong to a Poisson distribution with which the probability (in our case the blocking probability) of the limit value for the access control being exceeded can be estimated. The blocking probabilities and the limit values for access control are interrelated and can generally be converted into one another. With the method in accordance with the invention, for plurality of pairs formed from marginal nodes, the initialization step consists of setting the blocking probabilities by defining the limit values for the access control so that they are essentially the same. The initial blocking probabilities here are chosen so that they are big enough for no overload situations to occur in the network. This definition corresponds to fair handling of the data stream transmitted between the marginal nodes, to the extent that there is practically the same probability of it being allowed or rejected. With this definition however there is as yet no guarantee that the resources available to the network will be used efficiently. As regards efficient resource utilization, the method in accordance with the invention makes provision for lowering the blocking probability, i.e.

increasing the limit values for access control correspondingly until an overload situation occurs. The lowering of the blocking probabilities or the increasing of the limit values for the approval checking is undertaken such that the blocking probabilities remain essentially the same for the pair of marginal nodes. For the pairs of marginal nodes involved when the overload situation arises, the blocking probabilities are essentially set to the value at which the overload situation would be caused by raising the limit values. For example the blocking probabilities are lowered step-by-step and the value of the blocking probabilities and thereby also the value of the corresponding limits is then set for the pairs contributing to the overload situation to the value that it had in the step directly before the overload situation occurred.

The advantage of the invention is that in a network without explicit path reservation limits can be defined for access control in a balanced and resource-efficient way. In accordance with a development of the invention limit values are defined for all pairs of the plurality of pairs formed from marginal nodes. In this case, for pairs which were not involved in the occurrence of the first overload situation, the blocking probability is further lowered simultaneously for all remaining pairs until a second overload situation occurs. For the pairs involved in the occurrence of the overload situation the blocking probabilities or the limit values are essentially frozen or maintained at the value which they had on occurrence or shortly before the occurrence of the overload situation. This step is then repeated until such time as limit values have been defined for all pairs i.e. the blocking probability is simultaneously increased for the remaining pairs until an overload situation occurs in which, for the

pairs involved in the overload situation, the blocking probability is retained until such time as the blocking probability is in place for all pairs.

With this development there are two points to note:

1. The development leads to an assignment of blocking probabilities or limit values for all pairs since a blocking probability of zero for a pair would mean that one would be able to transmit an unlimited amount of traffic between the edge points of the pair without any overload situation arising, which is evidently not the case for real networks.
2. The blocking probability or limit values for all pairs of the plurality of pairs is defined so that a lowering of the blocking probability of any given pair from the plurality would lead to an overload situation. In this sense an optimum utilization of the resources available to the network is produced.

The plurality of pairs includes for example all pairs of ingress nodes and egress nodes. In this case complete control of the traffic entering the network and leaving the network again is provided, or of the traffic of the class of service which is subject to access control. The setting of the limits or the choice of the blocking probabilities guarantees that no overload situation occurs; As a consequence definitive statements can be made about the quality-of-service features. The setting of limits for access control then opens up the possibility of quality of service with simultaneously a best possible utilization of the resources available.

Communication networks have physical restrictions for the transmission capacity over the connection sections or links which are enclosed by the network. The maximum transmission capacity of the individual links defines an upper limit for the traffic able to be transmitted over the link in question. Frequently limits on the traffic volume over the individual links are set lower than the maximum physical capacity in order to provide spare capacity on the one hand and on the other hand to prevent faults occurring in the network. In the latter case the focus is frequently on the resilience of a network, i.e. the capability of ensuring problem-free transmission even with failures of network elements. For the communication network discussed above the limits for the traffic on the individual links can for example be selected so that the failure of one (or more) links does not lead to the physical limits for the capacity of the other links being exceeded, i.e. even if a link fails the traffic which was subject to access control can be managed. An overload situation in the sense of the method in accordance with the invention can then be defined as the fact that the defined limits for the traffic volume on this link could have been exceeded on a connection section or a link of the network. Checking the communication network for the possibility of an overload situation can be undertaken for example with the aid of a model for the load distribution within the network. A check is made for example using a simulation program as to whether there are links in the communication network for which a maximum traffic load within the framework of the defined limits would lead to the permitted traffic volume for the link being exceeded. Another slightly modified definition would be that the limits covering volume of traffic on the individual links would be likely to be exceeded with high predefinable

probability. The pairs of marginal nodes which contribute to an overload situation produced by a limit value being exceeded for the traffic on a link would then be those to which paths could be assigned which run via the link which is causing the overload situation. In traffic theory the term traffic pattern is usually used to designate the real (instantaneous) traffic present at the inputs of the network. Checking for an overload situation can then be undertaken by checking whether the traffic pattern allowed by the limits, taking into account the routings undertaken in the network, would then lead to an overload situation or not.

The object of the invention is explained in more detail below within the context of an exemplary embodiment which refers to a Figure.

The Figure shows a network made up of nodes and links. In this case the marginal nodes R1 to R10 are indicated by solid circles. The internal nodes are indicated by non-solid circles. Links are illustrated by connectors between nodes. For the network different types of peripheral conditions can be defined which guarantee approval control at the margin of the network. The type of peripheral conditions can for example be selected to depend on the topology of the network. The form of the peripheral conditions helps to decide on the blocking probabilities for which an overload situation occurs in accordance with the inventive method. Possible peripheral conditions are:

1. Limits for the traffic which is transmitted between two marginal nodes, i.e. a limit value in each case for a pair (Ri,Rj), $j,i \in \{1,..,10\}$, which is defined by two marginal nodes.

2. Limit values for all ingress and egress nodes. If we assume that all marginal nodes Ri, i ∈ {1,..,10} are both ingress and egress nodes, this would produce 20 limit values, with two limit values, an ingress value and an egress value being assigned to a node in each case. For a flow which is to be transmitted from the ingress node Ri to the egress node Rj a check would then be made on whether the node would exceed the ingress limit for Ri or the egress limit for Rj. Exceeding the limit would result in rejection.

3. Ingress and egress limit values as for Point 2. but for all links of the network. This means that for each link L one has two limits per marginal node in each case. For the transmission of a flow from node Ri to node Rj the ingress limits of Ri and the egress limits of Rj would be checked which relate to links over which the flow is to be transmitted.

To simplify matters the explanation belows assumes the form of limits described in 1. above. They are to form the basis of a traffic model which allows an average volume of traffic between two marginal nodes Ri and Rj to be determined. The average traffic between two marginal nodes Ri and Rj is designated for simplicity's sake as Vij. Likewise Gij is taken as the limit value for the traffic transmitted from the ingress node Ri to the egress node Rj. A flow to be transmitted from Ri to Rj is allowed if the aggregated traffic between Ri and Rj would not exceed the limit Gij. In this case j,i ∈ {1,..,10} always applies. The traffic model uses the average traffic values Vij and assumptions about the statistical fluctuations, which belong to a Poisson distribution for example. To initialize the method the limit values Gij are set so (low) that the same blocking probabilities exist for all pairs (Ri,Rj) and that in addition

no overload situation occurs. Checking for the occurrence of an overload situation can be undertaken for example by determining, for the maximum traffic load allowed by the limit values with the inclusion of the routings within the network, the traffic load of the individual links and comparing this with the limits or capacity of the links. In accordance with the invention the blocking probabilities are lowered by the same percentage ratio and the limit values Gij correspondingly increased. In this case, with the aid or the traffic model, for a set of reduced, similar values of the blocking probability a corresponding set of limit values Gij (analytical or numerical) is determined, which - within the framework of the traffic model - defines a non-approval with the reduced blocking probability for all pairs (Ri,Rj). A check for overload follows. If no overload occurs the blocking probability is further lowered simultaneously for all communication links. This can for example occur through step-by-step lowering by 10 of the initial value. In one step, for example the fifth a bottleneck or an overload situation occurs on the link L, i.e. the limits for the capacity on this link would be exceeded by the choice of limits at step 5. The pairs (R1,R2), (R2,R1), (R1,R3), (R3,R1), (R1,R4) and (R4,R1) contribute to this overload situation for example. For these pairs the limits Gij or the blocking probability are then set to their value at step 4. The method is then continued for the remaining pairs (Ri,Rj) until limits Gij are defined for all pairs (Ri,Rj).

Patent claims

1. Method for the setting limit values for limiting traffic in a communication network with access controls, with the communication network being formed by nodes and links, and at least for of a part of the traffic which is to be transmitted over the network, an access control being performed by using a limit value,

in which

- for a plurality of pairs formed by marginal nodes (Ri,Rj) the limit values for the access control are set such that

-- the probability of traffic which is subject to access control not being approved is essentially the same for the transmission between the two marginal nodes of each pair, and

-- overload situations do not occur,

- the relevant limit values are not applied until an overload situation occurs, with the cancelation being undertaken so that,

the probability of a non-approval of traffic remains essentially the same for individual pairs (Ri,Rj), and

- the limit values for access control for the transmission between pairs of marginal nodes, in which the traffic transmitted between the marginal nodes contributes to the overload situation, is determined as essentially the value at which the overload situation would be caused by raising the limit values.

2. Method in accordance with claim 1,

characterized in that,

until all limit values are set for access control for the transmission between the marginal nodes of the individual pairs (Ri,Rj)

- as yet unset limit values for the access control continue to be raised until an overload situation occurs, with the raising being undertaken such that the probability of a non-approval of traffic between pairs (Ri,Rj), for which no limit values have yet been set remains essentially the same, and

- the limit values for access control for the transmission between pairs (Ri,Rj) of marginal nodes in which traffic transmitted between the marginal nodes contributes to an overload situation is continued at essentially the value at which the overload situation was caused by raising the limit values.

3. Method in accordance with claim 1 or 2,

characterized in that

marginal nodes are specified by ingress nodes and egress nodes.

4. Method in accordance with one of the previous claims, characterized in that

the marginal nodes include nodes of the network representing sources or sinks of traffic.

5. Method in accordance with one of the previous claims, characterized in that

the plurality of the pairs formed by marginal nodes (Ri,Rj) include all pairs (Ri,Rj) consisting of an ingress node and an egress node in each case.

6. Method in accordance with one of the previous claims,

characterized in that

access checks are made for all the traffic of a class of service.

7. Method in accordance with one of the previous claims, characterized in that

the access checks relate to the approval or rejection of individual flows.

8. Method in accordance with one of the previous claims, characterized in that

an overload situation is produced when in a scenario of high traffic load, in which the limit values for the access controls are still adhered to, a threshold value is exceeded on a link for the traffic transmitted over the link.

9. Method according to claim 8,

characterized in that

threshold values for the traffic transmitted over the relevant link are assigned to the links of the communication network such that, if one or more of the links fails, the traffic allowed within the framework of the access controls does not represent any overload.

10. Network node with means for executing a method in accordance with one of the claims 1 to 9.